BOEING

The Number of Planar Trees

David W. Walkup
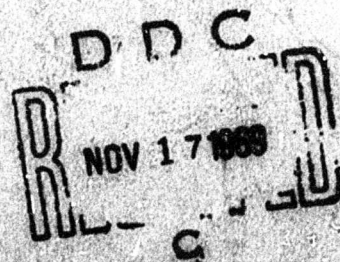
13

ABSTRACT

By a planar tree is meant a realization of a tree in the plane, and by an isomorphism between two planar trees is meant a mapping which is not only an isomorphism in the usual sense of trees but which also preserves the clockwise cyclic order of edges about each node. Explicit formulae are given for each of the following: (1) the number of nonisomorphic unrooted planar trees with n edges, (2) the number of nonisomorphic rooted planar trees with n edges, and (3) the number of nonisomorphic rooted planar trees with n edges such that the root is incident on exactly k edges, of which one is distinguished.

By a planar tree, rooted planar tree, etc., will be meant any realization of a tree, rooted tree, etc., in the plane. By an isomorphism between two planar trees, etc., will be meant any one-to-one mapping of the nodes and edges of one onto the other which is an isomorphism in the usual sense for such trees and which in addition preserves the clockwise cyclic order of the edges about each node. The results obtained in this note, contained in Theorems 1 through 3 below, extend results obtained by Harary, Prins, and Tutte [1]. In the interest of simplifying formulae, all the results given here are stated in terms of the number of edges of a tree, rather than the number of nodes as was done in [1].

THEOREM 1. Let $R(n)$ denote the number of nonisomorphic (undirected, unlabeled) rooted planar trees with n edges. Then for $n \geq 1$,

$$(1) \qquad R(n) = \frac{1}{2n} \sum_{s|n}{}' \varphi\left(\frac{n}{s}\right)\binom{2s}{s} \quad ,$$

where φ denotes the Euler function.

THEOREM 2. Let $F(n)$ denote the number of nonisomorphic (free, undirected, unlabeled) planar trees with n edges. Then for $n \geq 1$,

$$(2) \qquad F(n) = \frac{1}{2n(n+1)}\binom{2n}{n} + \frac{1}{4n}\left(\binom{n+1}{\frac{n+1}{2}}\right) + \frac{1}{n}\varphi(n)$$

$$+ \frac{1}{2n} \sum_{\substack{s|n \\ 1<s<n}} \varphi\left(\frac{n}{s}\right)\binom{2s}{s} \quad ,$$

where the second of the four terms on the right in (2) is to be under-
stood as zero if n is even.

Note that (2) has been written so that it reduces to the first
line if n is prime. In this form the asymptotic behavior of F(n)
is clear.

THEOREM 3. Let E(n,k) denote the number of nonisomorphic
rooted planar trees with n edges such that the root is incident on
exactly k edges, one of which is distinguished. Then

$$(3) \qquad E(n,k) = \frac{k}{2n-k} \binom{2n-k}{n-k} = \binom{2n-k}{n-k} - 2\binom{2n-k-1}{n-k-1}$$

In the terminology of Knuth [2; p. 306, p. 389] E(n,k) is the
number of nonequivalent ordered trees with n edges and k edges at
the root.

Proof of Theorem 3. With each tree T from the class enumerated
by E(n,k) we will associate a string s(T) of left and right pa-
rentheses. As a first step we convert T into an ordered tree
[2; p. 306] by redrawing it as in Fig. 1 with the root r(T) upper-
most and the distinguished edge e(T) to the left of the other edges
depended from the root. An examination of Fig. 1 together with the
following remarks should make clear how the string s(T) is derived
from the ordered tree T. To each of the n-k edges of T not
incident on the root there corresponds a matched pair of left and
right parentheses in s(T). Each node p_i, $1 \leq i \leq k$, adjacent to the
root determines a subtree T_i consisting of p_i and all nodes and

edges below it in T. To each such subtree there corresponds a (possibly empty) group of matched parentheses in s(T), and each such group is set off on the right by an unmatched right parenthesis. Thus the last character in the string s(T) is always an unmatched right parenthesis, whereas the first character is an unmatched right parenthesis if and only if the distinguished edge e(T) is a terminal edge. In all there are n-k left parentheses, n-k matched right parentheses, and k unmatched right parentheses in s(T), for a total of 2n-k characters. Let S(n,k) denote the set of *all* strings of length 2n-k consisting of n-k left parentheses and n right parentheses, and let S'(n,k) denote the subset consisting of all strings s(T) formed as T ranges over the class enumerated by E(n,k). Now observe that each string s in S'(n,k) determines the planar tree from which it is derived to within isomorphisms, and hence

$$(4) \qquad\qquad E(n,k) = |S'(n,k)| \ .$$

This follows easily from the observation that the matching of parentheses in s (and hence the ordered tree T) can be reconstructed unambiguously by first matching the innermost pairs of parentheses in any order and working outward, leaving exactly k unmatched right parentheses. This is essentially a standard result concerning the parsing of strings of parentheses. (If desired an extra k left parentheses can be imagined to the left of the string s.) However with a little extra thought it can be seen that the same procedure (i.e., working outward from the innermost pairs until exactly k unmatched parentheses remain) will

also work on *any* member of $S(n,k)$ provided it is first closed up
into a circular string. Clearly then each string in $S(n,k)$ can be
cyclically permuted to form a string in $S'(n,k)$, specifically by
placing any unmatched right parenthesis at the end. More particularly,
let s be any string in $S(n,k)$ and let C be the equivalence class
in $S(n,k)$ consisting of all strings which are cyclically equivalent
to s. Then $|C| = (2n-k)/q$, where $q \geq 1$ accounts for the fact that
s may be periodic. It is obvious that q must divide both $(2n-k)$
and k, and in fact $|C \cap S'(n,k)| = k/q$. Since the relation

$$\frac{|C \cap S'(n,k)|}{|C|} = \frac{k}{2n-k}$$

holds for each equivalence class, it must hold also for $S(n,k)$
itself, that is,

$$|S'(n,k)| = \frac{k}{2n-k} \binom{2n-k}{n-k} .$$

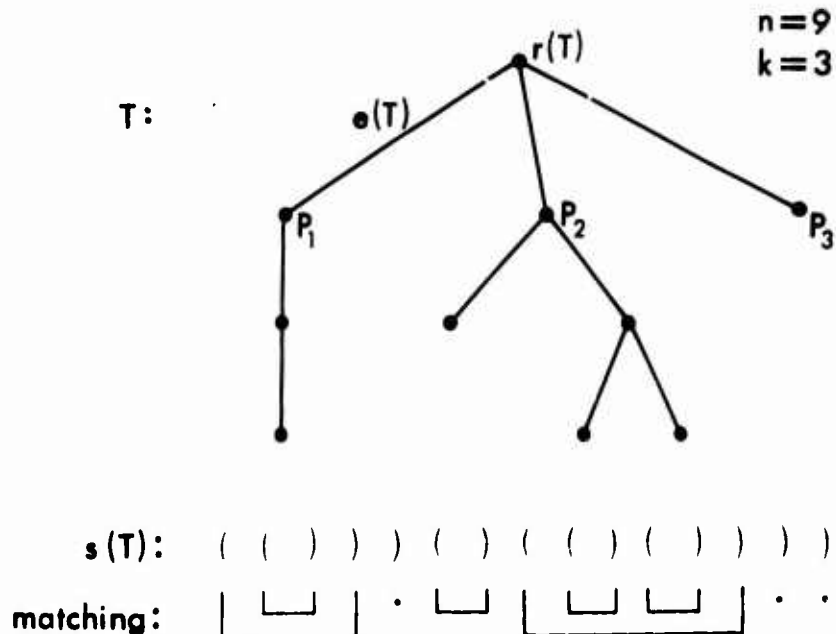The desired result (3) follows immediately from (4).

Fig. 1

Proof of Theorem 1. Consider the following five conditions which may be imposed separately or in combinations on a rooted planar tree T:

   (i)   T has n edges.

  (ii)   There are k edges incident on the root.

 (iii)   Exactly one edge of T incident on the root is distinguished.

  (iv)   The group of automorphisms of T which fix the root, but not necessarily the distinguished edge (if any), is of order q.

   (v)   The group of automorphisms of T which fix the root, but not necessarily the distinguished edge (if any), is of order a multiple of q.

Let $R(n,k,q)$, $E(n,k,q)$, and $E^*(n,k,q)$ denote the number of non-isomorphic rooted planar trees satisfying the following indicated combinations of the above conditions:

$$R(n) \ \ldots \ (i)$$
$$R(n,k,q) \ \ldots \ (i), \ (ii), \ (iv)$$
$$E(n,k) \ \ldots \ (i), \ (ii), \ (iii)$$
$$E(n,k,q) \ \ldots \ (i), \ (ii), \ (iii), \ (iv)$$
$$E^*(n,k,q) \ \ldots \ (i), \ (ii), \ (iii), \ (v)$$

For comparison the corresponding combinations of conditions defining $R(n)$ and $E(n,k)$ have been included. Note that each of the quantities $R(n,k,q)$, $E(n,k,q)$, and $E^*(n,k,q)$ is necessarily zero unless q divides both n and k. Thus clearly

(5)
$$R(n) = \sum_{q|n} \sum_{m=1}^{n/q} R(n,mq,q) \ .$$

Also,

(6)
$$R(n,mq,q) = \frac{1}{m} E(n,mq,q) \ ,$$

since for each tree in the class enumerated by $R(n,mq,q)$ there are only m isomorphically distinguishable edges at the root. Next, by the definitions of $E^*(n,k,q)$ and $E(n,k,q)$, $E^*(n,k,q) = \sum_{t \geq 1} E(n,k,tq)$, but because $E^*(n,k,q)$ and $E(n,k,q)$ are zero unless q divides k,

$$E^*(n,k,\frac{k}{m}) = \sum_{\ell|m} E(n,k,\frac{k}{\ell}) \quad \text{for} \quad m|k \ .$$

By the Moebius inversion formula,

$$E(n,k,\frac{k}{m}) = \sum_{\ell|m} \mu(\ell) \ E^*(n,k,\frac{\ell k}{m}), \quad \text{for} \quad m|k \ ,$$

or equivalently,

(7)
$$E(n,mq,q) = \sum_{\ell|m} \mu(\ell) E^*(n,mq,\ell q) \ .$$

Finally from the correspondence illustrated in Fig. 2 it is apparent that

(8)
$$E^*(n,i,j) = E(\frac{n}{j},\frac{i}{j}) \quad \text{provided} \quad j|n, \quad j|i \ .$$

Thus combining (5) through (8) we have

$$R(n) = \sum_{q|n} \sum_{m=1}^{n/q} \sum_{\substack{\ell|m \\ \ell|\frac{n}{q}}} \mu(\ell) \ \frac{1}{m} \ E(\frac{n}{\ell q},\frac{m}{\ell}) \ ,$$

where the condition $\ell \mid (n/q)$ reflects the proviso in (8). With a change of variables $m = \ell t$ and $q = n/r$ and a rearrangement of the summation, this becomes

$$R(n) = \sum_{r \mid n} \sum_{\ell \mid r} \sum_{t=1}^{r/\ell} \mu(\ell) \frac{1}{\ell t} E(\frac{r}{\ell}, t)$$

A further change of variable $r = s\ell$ and rearrangement yield

(9) $$R(n) = \sum_{s \mid n} \left[\sum_{\ell \mid \frac{n}{s}} \frac{1}{\ell} \mu(\ell) \right] \left[\sum_{t=1}^{s} \frac{1}{t} E(s,t) \right] .$$
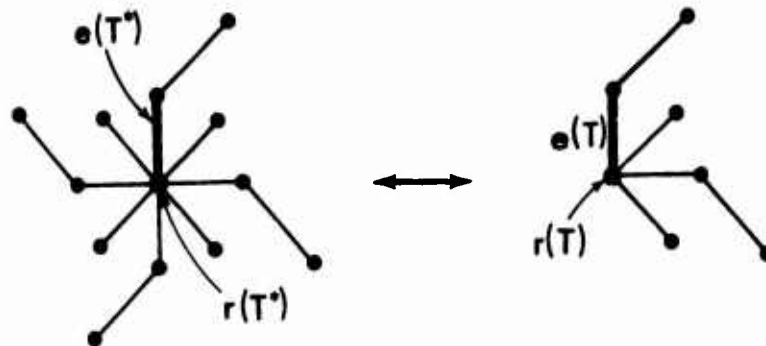
The first bracketed sum can be simplified using the elementary identity

$$\sum_{\ell \mid w} \frac{1}{\ell} \mu(\ell) = \frac{1}{w} \varphi(w) .$$

The second bracketed sum in (9) can be simplified using (3) and some elementary binomial identities [2, p.53] as follows:

$$\sum_{t=1}^{s} \frac{1}{t} E(s,t) = \sum_{t=1}^{s} \frac{1}{2s-t} \binom{2s-t}{s}$$

$$= \frac{1}{s} \sum_{t=1}^{s} \binom{2s-t-1}{s-1}$$

$$= \frac{1}{s} \binom{2s-1}{s}$$

$$= \frac{1}{2s} \binom{2s}{s} .$$

The desired result (1) is now immediate.

$$E^*(12,8,2) \;=\; E(6,4)$$

Fig. 2

Proof of Theorem 2. Let $E_k(x)$ denote the enumerating function

$$E_k(x) \;=\; \sum_{n \geq k} E(n,k)\, x^n .$$

From the definition of $E(n,k)$ in terms of trees it follows immediately that

(10) $$E_k(x) \;=\; [E_1(x)]^k .$$

From Theorem 3 of [1] we have

$$t(x) \;=\; T(x) \;-\; \frac{1}{2x^2}\, [P^2(x) - P(x^2)]$$

where $t(x)/x$ is the enumerating function for $F(n)$, $T(x)/x$ is the enumerating function for $R(n)$, and $P(x)/x = E_1(x)$. (The common factor $1/x$ accomplishes the translation from enumeration by nodes to enumeration by edges.) In view of (10) then,

$$F(n) = R(n) - \frac{1}{2} E(n+1,2) + \frac{1}{2} E(\frac{n+1}{2},1) \ ,$$

where the final term is to be interpreted as zero is n is even.
Applications of (1), (3), and elementary binomial identities yield
the desired result (2).

$$F(n) = R(n) - \frac{1}{2} E(n+1,2) + \frac{1}{2} E(\frac{n+1}{2},1) \ \bullet$$

REFERENCES

1. F. Harary, G. Prins, and W. T. Tutte, The number of plane trees,
   Kon. Ned. Akad. van Wet., Proc. Ser. A, <u>67</u> (1964), 319-329 (=
   Indagationes Math., <u>26</u>).

2. D. E. Knuth, The Art of Computer Programming, Vol. I, Addison-
   Wesley, Reading, Massachusetts, 1968.